УДК 519.237.5: 621.9

Лапач С.М.
НТУУ «Київський політехнічний інститут», м. Київ, Україна

ПРОБЛЕМИ ПОБУДОВИ РЕГРЕСІЙНИХ МОДЕЛЕЙ ПРОЦЕСІВ РІЗАННЯ МЕТАЛІВ

Lapach S.
National Technical University of Ukraine «Kyiv Polytechnic Institute», Kyiv, Ukraine (mmi@kpi.ua)

PROBLEMS CONSTRUCTION REGRESSIAN MODEL OF CUTING PROCESS

Виконано порівняння різних підходів до специфікації моделі процесу різання і їх обґрунтованості, як з предметної галузі, так і з боку математичної статистики. На прикладах розв'язання конкретних задач показані недоліки традиційної мультиплікативної моделі зі статистичної і фізичної точок зору, а також протиріччя між статистичною оцінкою адекватності в регресійному аналізі і вимогами до моделі, які випливають з предметної галузі. Пропонується не використовувати мультиплікативні моделі в зв'язку з відсутністю їх обґрунтування і як такі, які приводять до зміни вихідної задачі. Адекватність моделі пропонується оцінювати виходячи з вимог предметної галузі.
Ключові слова: лінійний регресійних аналіз, моделювання процесів різання, специфікація моделі, адекватність моделі, перетворення вихідних факторів, поліноми Чебишева.

Вступ

Проектування технологічних процесів – це вибір і побудова систем зв'язків, дія яких приводить до бажаного кінцевого результату: продукції з потрібними характеристиками якості. При цьому передаточні моделі мають випадкову природу і будь-яка технологічна система може розглядатись тільки як «чорний ящик» [1]. Тому при побудові таких залежностей використовується регресійний аналіз. Поширення верстатів з числовим програмним керуванням та систем автоматичного проектування вимагають інформаційної технології їх побудови. Не зважаючи на таку потребу, в галузі обробки різанням остання ґрунтовна спеціальна робота по побудові таких залежностей написана досить давно [2]. За цей час теорія планування експериментів і регресійний аналіз набули значного розвитку, досягнення яких в цілому (за виключенням робіт, пов'язаних зі школою Радченко С.Г.) не набули широкого застосування в прикладних дослідженнях по обробці різанням металів. Крім того, в рамках застосування регресійного аналізу для опису процесів різання залишаються нерозв'язаними ряд принципових питань стосовно специфікації моделі, як загальної, так і часткової. В статті розглядається вплив вибору загальної форми рівняння (мультиплікативна – адитивна) на можливість визначення статистично значимих регресорів, а також неоднозначність статистичного критерію адекватності по відношенню до проблеми визначення часткової структури.

Мета

Досліджується формування специфікації рівняння регресії при побудові емпіричних моделей процесів різання. Виконується аналіз існуючих підходів, їх порівняння і формування рекомендацій по побудові, яке б відповідало особливостям предметної галузі.

Дослідження

Специфікацією математичної моделі називається вибір форми функції, яка буде використовуватися для опису процесу. Ідентифікацією – отримання оцінок параметрів цієї функції (в даному випадку коефіцієнтів регресії). У випадку лінійної регресії може складатися з двох етапів: вибір загальної форми і встановлення конкретного переліку цих елементів. Це викликано неможливістю встановити конкретний перелік елементів моделі через перевірку статистичної значимості коефіцієнтів, виключаючи випадок повного факторного експерименту. Існує два основних підходи до визначення специфікації рівняння лінійної регресії: поліноміальна і степенева залежність.

Метою роботи є проведення аналізу обох підходів з точки зору регресійного аналізу і вироблення рекомендацій по формуванню специфікації. Для наочності зробимо аналіз на прикладі реальних задач. Використаємо вихідні данні з книги [2]. Книга відповідає тогочасному рівню розвитку і застосування планування експерименту та регресійного аналізу.

При побудові степеневої залежності в викладі [1] маємо наступні кроки:

1. завдання загального виду функції (специфікація моделі);

2. лінеаризація (перетворення вихідних змінних);
3. оцінка коефіцієнтів регресії.
4. оцінка значимості коефіцієнтів регресії (формування часткової структури рівняння регресії);
5. оцінка адекватності.

Побудова лінійної регресії проходить наступні етапи (спрощено) [4, 5, 6]:

1. завдання загального виду функції (специфікація моделі);
2. побудова поліномів Чебишева і взаємодій (перетворення вихідних змінних);
3. формування часткової структури рівняння регресії;
4. оцінка коефіцієнтів регресії;
5. оцінка адекватності суміщена з уточненням структури;

Алгоритми регресійного аналізу детально викладені у вказаних джерелах і в даній роботі не приводяться.

Приклад 1.

Розглядається побудова залежності стійкості від режимів різання різців с пластинками зі сплаву Т14К8.

Опис задачі і послідовність побудови моделі взятий з [2, С.152]. Позначення в формулах (1) – (3) приведено за вказаним джерелом.

Рівняння постулюється у формі

$$T = Cv^x s^y t^z , \qquad (1)$$

де Т – показник стійкості в хв., С – постійний коефіцієнт; v – швидкість різання в м/хв.; s – подача в мм/оберт; t – глибина різання в мм. Для знаходження коефіцієнтів виконується лінеаризація логарифмуванням, яка приводить рівняння до виду

$$M = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_{13} \qquad (2)$$

Тут М – логарифмоване значення стійкості; x_1 , x_2 , x_3 – логарифми v, s і t відповідно. Для перетворення змінних від натуральних значень до кодованих застосовується формула

$$x_i = \frac{2(\ln \tilde{x}_i - \ln \tilde{x}_{i\max})}{\ln \tilde{x}_{i\max} - \ln \tilde{x}_{i\min}} + 1 \qquad (3)$$

Сам автор вказує, що логарифмування приводить до зміщення оцінок параметрів і порушення передумов про адитивність входження помилок в вихідну модель.

Використовувався план повного факторного експерименту виду 2^3//8. ПФЕ є ідеальним варіантом з точки зору планування експерименту. Значення рівнів факторів приведено в табл. 1.

Таблиця 1

Значення рівнів факторів прикладу 1

Рівень фактора	Фактор		
	v, м/хв	s, мм/об	t, мм
Верхній (+)	226	0,2	2,0
Нижній (–)	56	0,049	0,5

Приведена модель

$$\hat{Y} = 4,04 - 0,96x_1 - 0,29x_2 + 0,13x_3 \qquad (4)$$

Перевірки адекватності, інформативності і стійкості відсутні. Це досить поширена ситуація для того часу в описі прикладних задач. Структура моделі задана (теж звичайно), а значимість коефіцієнтів отриманої моделі не перевіряється. Сказано тільки, що квадратичні члени відсутні. Значимість коефіцієнтів зазвичай перевірялась. У цього ж автора в інших задачах вона перевіряється. Причини цього, на нашу думку, ґрунтуються на протиріччі структури постульованої моделі з тією, яка має бути, виходячи з прийнятою за статистичними показниками (після перевірки значимості), що буде надалі показано..

Виконаємо перерахунок моделі за тими ж умовами традиційними методами. Розрахунки виконувались за допомогою стандартної функції табличного редактора Excel [5] (перераховані тільки значення критеріїв Стьюдента і середньоквадратична похибка коефіцієнта регресії в зв'язку з наявністю дисперсії відтворюваності).

Отримана модель

$$\hat{Y} = 4,04407 - 0,95988x_1 - 0,29018x_2 + 0,128627x_3 \qquad (5)$$

Як видно, значення коефіцієнтів членів моделі практично співпадають з (4).

Модель інформативна. Коефіцієнт множинної кореляції 0,950. Цей коефіцієнт статистично значущий: F_R=12,36 > $F_{кр}$= 4,12 при рівні значущості α=0,05 та степенях свободи ν_1 = 4 і ν_2 = 7. Модель пояснює 90,3% загального розсіювання. Середня похибка апроксимації 8,02%. Значення критерію для перевірки значущості коефіцієнтів регресії Стюдента t=2,12.

В табл.2 представлені перевірка значимості коефіцієнтів за критерієм Стюдента.

Таблиця 2

Статистичні характеристики коефіцієнтів регресії для моделі головних ефектів

Назва регресора	Коефіцієнти	Середньоквадратична похибка	t-статистика
b_0	4,04407	0,060038	67,36
X_1	-0,95988	0,060038	-16,00
X_2	-0,29018	0,060038	-4,83
X_3	0,128627	0,060038	2,14

Після цього було виконано пошук структури моделі за цими ж даними, використовуючи оцінку значимості коефіцієнтів регресії за критерієм Стюдента. Результати приведені в табл. 3.

Таблиця 3

Статистичні характеристики коефіцієнтів регресії для моделі зі взаємодіями

Назва регресора	Коефіцієнти	Середньоквадратична похибка	t-статистика
b_0	4,04407	0,060038	67,36
X_1	-0,95988	0,060038	-16,00
X_2	-0,29018	0,060038	-4,83
X_3	0,128627	0,060038	2,14
$X_1\ X_2$	-0,2324	0,060038	-3,87
$X_2\ X_3$	-0,15981	0,060038	-2,66

В модель при рівні значущості α=0,05 входять дві взаємодії і всі три головних фактори.

$$\hat{Y} = 4,04407 - 0,95988x_1 - 0,29018x_2 + 0,128627x_3 - 0,2324x_1x_2 - 0,15981x_2x_3 \qquad (6)$$

Модель інформативна. Коефіцієнт множинної кореляції 0,986. Цей коефіцієнт статистично значущий: F_R=14,35> $F_{кр}$= 4,74 при рівні значущості α=0,05 та степенях свободи ν_1 = 2 і ν_2 = 7. Модель пояснює 90,5% загального розсіювання. Середня похибка апроксимації 4,64%.

Як бачимо, навіть в рамках тогочасного підходу модель повинна мати іншу структуру (порівнюючи з (5)).

Перерахуємо модель за ПРІАМ (алгоритм покрокового визначення структури) [3, 4, 5]. Отримана модель за структурою не співпадає з (5) і співпадає з (6):

Дисперсії дослідів однорідні $G_{розр}$=0,25 < $G_{кр}$= 0,99 при рівні значущості α=0,05 та степенях свободи ν_1 = 8 і ν_2 = 2. Модель адекватна ($F_{розр}$=2,04<$F_{кр}$=3,63) при рівні значущості α=0,05 та степенях свободи ν_1 = 5 і ν_2 = 2. Модель інформативна. Коефіцієнт множинної кореляції 0,986. Цей коефіцієнт статистично значущий: F_R=14,36> $F_{кр}$= 2,77 при рівні значущості α=0,05 та степенях свободи ν_1 = 5 і ν_2 = 18. Інформативність моделі задовільна: для критерію Бокса-Веца γ=1. Модель обчислювально стійка: число обумовленості cond(X^TX) = 1. Модель структурно стійка: матриця, за якою розраховані коефіцієнти регресії ортогональна.

Модель пояснює 97,3% загального розсіювання. Середня похибка апроксимації 4,27%.

Впливом x_1 пояснюється 81,37% загального розсіяння, а x_2 – 7,43%, x_3 – 1,46%., взаємодією x_1x_2 – 4,76%, взаємодією x_2x_3 – 2,25%.

Другий варіант загального завдання специфікації – використання лінійної регресії в формі (7)

$$\hat{y} = b_0 + \sum_{i=1}^{k} b_i f_i(x_1,...x_m) \qquad (7)$$

Тут $f_i(x_1,...x_m)$ – довільна функція від вихідних факторів. В загальному випадку, при відсутності інформації чи підстав для прийняття іншого рішення, в якості $f_i(x_1,...x_m)$ виступають поліноми Чебишева (головні ефекти та взаємодії).

Виконаємо перерахунок у відповідності з сучасними вимогами без логарифмування, але з перетворенням вихідних даних до поліномів Чебишева. Розрахунки виконано за допомогою програмного засобу ПРІАМ[3].

Ситуація виявляється ще більш складна, ніж здавалась до цього.

Адекватною є наступна модель (критичне значення критерія Стьюдента $t_{0,05;16}$=2,11):

$$\begin{aligned} &\hat{y} = 91,4583 - 64,4583x_1 + 19,5417x_3 - 14,375x_1x_3, \\ &\textit{де} \\ &x_1 = 0,0117647(X_1 - 141); \\ &x_3 = 1,333333(X_3 - 1,25). \end{aligned} \qquad (8)$$

Дисперсії дослідів однорідні $G_{розр}$=0,45 < $G_{кр}$= 0,99 при рівні значущості α=0,05 та степенях свободи ν_1=8 і ν_2 = 2. Модель адекватна ($F_{розр}$=2,15<$F_{кр}$=5,84) при рівні значущості α=0,05 та степенях свободи ν_1 = 4 і ν_2 = 4. Модель інформативна. Коефіцієнт множинної кореляції 0,981. Цей коефіцієнт статистично значущий:

F_R=34,42>$F_{кр}$= 3,10 при рівні значущості α=0,05 та степенях свободи ν_1 = 3 і ν_2 = 20. Інформативність моделі задовільна: для критерію Бокса-Веца γ=1. Модель обчислювально і структурно стійка.

Модель пояснює 96,27% загального розсіювання. Середня похибка апроксимації 41,18%.

Впливом x_1 пояснюється 84,32% загального розсіяння, x_3 – 7,75%., взаємодією x_1x_3 – 4,19%.

Не зважаючи на те, що для формування структури використовується, як і попередньо, критерій Стьюдента, структура моделі відрізняється.

Оскільки модель має велику похибку побудуємо більш точну модель. При цьому використаємо середні значення відгуків для можливості не враховувати статистичний критерій адекватності для обмеження введення членів в моделі повного факторного експерименту.

$$\hat{y} = 91,4583 - 64,4583x_1 + 19,5417x_3 - 14,375x_1x_2 - 11,5417x_2 - 6,45833x_2x_3$$

$$де$$

$$x_1 = 0,0117647(X_1 - 141); \tag{9}$$

$$x_2 = 13,245(X_2 - 0,1245);$$

$$x_3 = 1,333333(X_3 - 1,25).$$

Модель інформативна. Коефіцієнт множинної кореляції 0,999. Цей коефіцієнт статистично значущий: F_R=222,28> $F_{кр}$= 19,30 при рівні значущості α=0,05 та степенях свободи ν_1 = 5 і ν_2 = 2. Інформативність моделі добра: для критерію Бокса-Веца γ=3. Модель обчислювально стійка: число обумовленості COND = 1. Модель структурно стійка: матриця кореляції регресорів ортогональна.

А ось з адекватністю ми бачимо протиріччя між статистичними критеріями. Якщо вважати, що дисперсія відтворюваності нам невідома, то модель ($F_{розр}$=19,30>$F_{кр}$=14,07; перевірки при відсутності повторних дослідів [7]) при рівні значущості α=0,05 та степенях свободи ν_1 = 5 і ν_2 = 1. А оскільки вона відома, то модель виявляється неадекватною ($F_{розр}$=22,39>$F_{кр}$=4,82) при рівні значущості α=0,05 та степенях свободи ν_1 = 8 і ν_2=5.

Модель пояснює 99,82% загального розсіювання. Середня похибка апроксимації 7,54%.

Впливом x_1 пояснюється 84,33% загального розсіяння, x_2 – 2,70%, x_3 – 7,75%., взаємодією x_1x_3 – 4,19 взаємодією x_2x_3 – 0,85%.

Як ми бачимо, структура моделі не співпадає повністю з (7) і навіть сила впливу для частини регресорів впорядкована по іншому.

Якщо виконувати всі дії у відповідності з тогочасними вимогами, то буде отримана зовсім інша модель, а не та, яка постульована. Як з тексту цього підрозділу, так і з усієї книги ясно, що автору все це відомо, але він взяв модель у вигляді, якому її традиційно представляють. Як ми бачимо, таке представлення не відповідає фактичному положенню. ПРІАМ дозволяє отримати цю модель автоматично.

Прийняті перетворення приводять до порушення передумов регресійного аналізу, деформації структури моделі, зміщення оцінок параметрів. Фізичного обґрунтування вибору такої структури не існує.

Побудова моделі без логарифмування дає іншу структуру моделі навіть при розрахунку за ідентичними алгоритмами і програмами (формули (8) і (9)). Для розуміння причин такої сильної невідповідності див. табл. 4.

Таблиця 4

Зміна коефіцієнтів кореляції з відгуком при перетвореннях вихідних змінних

Перетворення	Коефіцієнти кореляції з відгуком		
	x_1	x_2	x_3
Для натуральних змінних	0,999961	0,000888	0,008823
Для поліномів Чебишева	0,999961	0,000888	0,008823
Для трансформованих за Кацевим	-0,88637	-0,26232	0,127788

З табл. 4 зрозуміло, перетворення з логарифмуванням некоректні (на відміну від поліномів Чебишева), оскільки змінюють співвідношення між коефіцієнтами кореляції. Останнє приводить до іншої структури моделі.

Звернемо увагу, що, крім вказаного, існують і проблеми з застосуванням суто статистичних показників при формуванні моделі. План повного факторного експерименту дозволяє нам розрахувати відразу всі коефіцієнти і їх статистичні характеристики для прийняття рішення про часткову структуру моделі.

В таблиці 5 приведені структури ряду моделей і їх основні характеристики. Ці характеристики, як статистичні, так і ні, дозволяють отримати уявлення про якість моделі.

Таблиця включає наступні варіанти моделей (в порядку розміщення): головних ефектів (1); ефектів, для яких розрахункове значення критерію Стьюдента більше одиниці (2); статистично значимі ефекти (3); головні ефекти плюс статистично значима взаємодія (4); головні ефекти плюс всі парні взаємодії (5); один головний ефект, який пояснює біля 90% розсіяння (6); найбільш сильні ефекти (7). Для вибору будь якого варіанту можливо найти обґрунтування в відповідній літературі.

Якщо ми подивимось таблицю 5, то побачимо, що різні характеристики входять в протиріччя один до одного. Це при цьому, що ми маємо справу з ідеальним варіантом – матрицею повного факторного експерименту.

Наприклад, найвищу інформативність (F_R) має модель (6), при цьому вона не адекватна і враховує тільки одну з трьох незалежних змінних. Найбільш повно описує дані (R^2) модель (5), але при цьому вона має найнижчу інформативність. Модель, яка має найвищу інформативність серед адекватних (3), включає тільки дві з трьох незалежних змінних. Зауважимо також, що із семи моделей чотири адекватні, тобто з точки зору математичної статистики рівнозначні. Навряд чи ми можемо вважати їх рівнозначними з точки зору технології машинобудування: вони навіть включають різні незалежні змінні.

Таким чином, як різні статистичні характеристики, так і вимоги предметної галузі входять в протиріччя при формуванні структурі і визначення адекватності навіть в найпростіших ситуаціях.

Таблиця 5

Порівняння статистичних характеристик різних за структурою моделей за прикладом 1

Назва регресора	Розрахункове значення критерію Стьюдента	Структура моделі (варіанти)						
		1	2	3	4	5	6	7
b_0	39,89	X	X	X	X	X	X	X
x_1	-9,47	X	X	X	X	X	X	X
x_2	-2,86	X	X	X	X	X		X
x_3	1,27	X	X		X	X		
x_1x_2	-2,29		X	X	X	X		X
x_1x_3	-0,92					X		
x_2x_3	-1,58		X			X		X
Статистичні характеристики								
Частка загального розсіяння, пояснювана моделлю, R^2		0,90268	0,9729	0,9357	0,9503	0,9805	0,8137	0,9583
Розрахункове значення критерію Фішера для перевірки значимості коефіцієнта множинної кореляції, F_R		12,36	14,36	19,41	14,35	8,39	26,20	17,23
Частка врахованих незалежних змінних в моделі, %		100	100	66,67	100	100	33,33	100
Залишкова дисперсія		0,220477	0,122758	0,145548	0,149944611	0,176401	0,28132	0,125959
Розрахункове значення критерію Фішера для адекватності		3,67	2,04	2,42	2,50	2,94	4,69	2,10
Табличне значення критерію Фішера для адекватності		3,01	3,63	3,01	3,01	4,49	2,74	3,24
Прийняття гіпотези про адекватність моделі		-	+	+	+	+	-	+

Приклад 2 (взято з [2, с.155).

Виконувалось дослідження залежності питомого зносу при плоскому шліфуванні зразків перерізом 18×16мм^2 зі сталей Р12 кругами з ельбору АЧК150×10×32×3 Л10Б1–100%.

Модель постулювалась у вигляді (1). Використовувався план повного факторного експерименту виду $2^3//8$. Натуральні значення рівнів факторів приведені в табл. 6.

Таблиця 6

Значення рівнів факторів для прикладу 2

Рівень фактора	Фактор		
	v, м/с	s, м/хв	t, мм/дв.ход
Верхній (+)	30	5,0	0,04
Нижній (–)	15	1,5	0,01

При перетвореннях аналогічних описаним раніше ((2), (3)) отримується наступна модель для стійкості в метрах шляху шліфуванні

$$\hat{y} = 1,161 - 0,112x_1 + 0,377x_2 + 0,722x_3 \qquad (10)$$

В роботі перевіряється значимість коефіцієнтів моделі, включаючи подвійні взаємодії і перевірки адекватності. Перевірка адекватності спрощена ($S^2_{зал} < S^2_{відт}$). З таблиці 7 видно, що логарифмування деформує структуру зв'язків між відгуком і незалежними змінними, що не дозволяє отримати модель, яка відповідає експериментальним даним. Звернемо увагу, що для натуральної моделі значимими будуть ефекти x_2, x_3, x_2x_3, а для логарифмованої – тільки x_2 і x_3, з деякою натяжкою, можливо, x_1. Причому, на відміну від прикладу 1, спотворення зв'язків тут більш відчутне і стосується вже навіть головних ефектів. Не зупиняючись більше на цьому факті, розглянемо проблеми формування структури і лінійній моделі.

Таблиця 7

Деякі статистичні показники для дослідження значимості прикладу 2

Назва регресора	Коеф. Кореляції з відгуком		Розрахункове значення t	
	Натуральна	Логарифмована	Натуральна	Логарифмована
b_0	–	–	8,48	14,73
x_1	-0,14331	-0,13186	-0,95	-1,36
x_2	0,800645	0,878574	5,31	9,07
x_3	0,477308	0,457468	3,17	4,73
x_1x_2	-0,09475	-0,01416	-0,63	-0,15
x_1x_3	-0,04382	0,000174	-0,29	0,002
x_2x_3	0,315047	0,030605	2,09	0,32

ПРІАМ в автоматичному режимі будує наступну лінійну регресійну модель

$$\begin{aligned} &\hat{y} = 4,49583 + 2,81667x_2 + 1,67917x_3 + 1,10833x_2x_3, \\ &де \\ &x_2 = 0,571429(X_2 - 3,25); \\ &x_3 = 66,6667(X_3 - 0,025). \end{aligned} \qquad (11)$$

Дисперсії дослідів однорідні $G_{розр}=0,47 < G_{кр}= 0,99$ при рівні значущості $\alpha=0,05$ та степенях свободи $\nu_1=8$ і $\nu_2 = 2$. Модель адекватна ($F_{розр}=2,85<F_{кр}=5,84$) при рівні значущості $\alpha=0,05$ та степенях свободи $\nu_1 = 3$ і $\nu_2 = 4$. Модель інформативна. Коефіцієнт множинної кореляції 0,984. Цей коефіцієнт статистично значущий: $F_R=40,48 > F_{кр}= 3,10$ при рівні значущості $\alpha=0,05$ та степенях свободи $\nu_1 = 3$ і $\nu_2 = 20$. Інформативність моделі задовільна: для критерію Бокса-Веца $\gamma=2$. Модель обчислювально і структурно стійка.

Модель пояснює 96,8% загального розсіювання. Середня похибка апроксимації 10,99%.

Користуючись табл. 8 (аналогічної табл. 5) розглянемо проблеми з застосуванням статистичних показників при формуванні моделі, користуючись властивостями плану повного факторного експерименту, який дозволяє нам розрахувати відразу всі коефіцієнти і їх статистичні характеристики для прийняття рішення про часткову структуру моделі.

Таблиця включає наступні варіанти моделей (в порядку розміщення): головні ефекти плюс всі парні взаємодії (1); головних ефектів (2); статистично значимі ефекти (3);

ефектів, для яких розрахункове значення критерію Стьюдента більше одиниці (2); головні ефекти плюс статистично значима взаємодія (4).

Найвищу інформативність (F_R) і найбільш повно описує дані (R^2) модель (1), при цьому вона не адекватна. і враховує тільки одну з трьох незалежних змінних. З адекватних моделей найбільш повно описує дані і має найбільшу інформативність модель (4), але при цьому вона включає статистично не значимий член (x_1). І знову ж маємо три адекватні моделі з різною структурою.

Зауважимо, що випадку, коли матриця плану не є матрицею повного факторного експерименту (а така ситуація наявна в більшості реальних задач), то ситуація набагато складніша в зв'язку з закорельованістю факторів.

Таблиця 8

Статистичні характеристики моделей різної структури для прикладу 2

Назва регресора	Структура моделі (варіанти)			
	1	2	3	4
b_0	X	X	X	X
x_1	X	X		X
x_2	X	X	X	X
x_3	X	X	X	X

Продовження таблиці 8

x_1x_2	X			
x_1x_3	X			
x_2x_3	X		X	X
Статистичні характеристики				
Частка загального розсіяння, пояснювана моделлю, R^2	0,999693	0,889393	0,968109	0,988647
Частка врахованих незалежних змінних в моделі, %	100	29,35	29,35	65,31
Залишкова дисперсія	0,001678	2,737812	0,789375	0,374676
Розрахункове значення критерію Фішера для адекватності	1339,333	1,218385	2,846661	5,997404
Табличне значення критерію Фішера для адекватності	246,4639	3,006917	5,844117	8,692286
Прийняття гіпотези про адекватність моделі	–	+	+	+

Висновки

В даній роботі встановлено і на прикладах показані наступні факти.

1) Використання мультиплікативної моделі з логарифмічними перетворення для лінеаризації приводить до деформації статистичних зв'язків між незалежними змінними і відгуком, тобто порушується відповідність між елементами структури вихідної і перетвореної моделі.

2) При використання логарифмічних перетворень змінюється коефіцієнт кореляції Пірсона, а в регресійній моделі – розрахункове значення критерію Стьюдента. Це приводить до зміни структури моделі. При використання перетворень в вигляді поліномів Чебишева деформації статистичних зв'язків не відбувається.

3) Логарифмічні перетворення ведуть до зміщення оцінок коефіцієнтів регресії і порушення передумови регресійного аналізу про адитивність випадкової помилки.

4) Статистичні характеристики адекватності, інформативності і тісноти зв'язку моделі навіть для ідеального випадку повного факторного експерименту можуть входити в протиріччя одне з одним, а також можуть протирічити вимогам предметної галузі;

5) Існує множина адекватних моделей (формально рівнозначних) з різними структурами.

6) Статистичний критерій адекватності може давати протилежні відповіді в тій же задачі при різних умовах застосування.

В зв'язку з цим рекомендується не використовувати мультиплікативну форму рівняння регресії для опису процесів різання, а застосовувати адитивну форму лінійної регресії. Крім відсутності вказаних недоліків, ця форма має перевагу простоти інтерпретації і можливість враховувати практично будь яку форму залежності відгуку від окремих факторів. Це забезпечується тим, що лінійна регресія може бути алгебраїчною сумою довільних функцій.

Що стосується прийняття рішення про адекватність моделі, то доцільним є урахування статистичних характеристик тільки як дорадчих. В основу прийняття рішення прийняти загальний підхід математичного моделювання: модель адекватна, якщо вона відповідає сукупності вимог, сформульованих в предметній галузі до неї.

Перспективи подальших розвідок у цьому напрямку.

Формулювання сукупності стандартних вимог до математичної моделі, яка описує процеси різання, виходячи з технологій машинобудування і теорії різання металів, яка дозволить визначати її адекватність.

Аннотация. *Проведено сравнение различных подходов к спецификации модели процессов резания и их обоснованность со статистической и предметной точек зрения. На решении конкретных задач показаны недостатки традиционной мультипликативной модели (со статистической и физической точек зрения), а также противоречие между статистической оценкой адекватности в регрессионном анализе и требованиями к модели в предметной области. Предлагается не использовать мультипликативные модели в связи с отсутствием их обоснованности и как приводящие к*

изменению исходной задачи. Адекватность модели предлагается устанавливать исходя из требований предметной области.
Ключевые слова: линейный регрессионный анализ, моделирование процессов резания, спецификация модели, адекватность модели, преобразование исходных факторов, полиномы Чебышева

Abstract. *The comparison of different approaches to cutting operations model specification and their feasibility from statistical and substantial viewpoints was made. It is shown that the linearization to be used for solving the problem for the multiplicative model leads to a change in the correlation coefficients between the covariates and the response. The result is a distortion of the model structure (composition and relative importance of the covariates). When using Chebyshev polynomials distortion does not occur. Demonstrated a conflict between the various statistical indicators, which also ceteris paribus leads to different structure models. The solutions of specific tasks show the disadvantages of traditional multiplicative model (from statistical and physical viewpoints), also discrepancy between statistical estimate of adequacy in regression analysis and requirements to the model in object region. It is suggested not to use multiplicative models due to absence of their feasibility and as leading to changes of original problem. The adequacy of model is suggested to be estimated coming from the requirements of object region.*
Keywords: linear regression analysis, cutting process modeling, model specification, model adequacy, transformation of background factors, Chebyshev polynomial

Бібліографічний список використаної літератури

1. *Колесов И.М.* Основы технологии машиностроения / Колесов И.М. Основы технологии машиностроения: Учеб. для машиностроит. спец. вузов. –М.: Высш. шк.., 2001. –591с.
2. *Кацев П.Г.* Статистические методы исследования режущего инструмента / Кацев П.Г. Статистические методы исследования режущего инструмента. Изд. 2-е, перераб. и доп. –М.: Машиностроение, 1974. –231с.
3. *Планирование,* регрессия и анализ моделей PRIAM (ПРИАМ) / Лапач С.Н., Радченко С.Г., Бабич П.Н. Планирование, регрессия и анализ моделей PRIAM (ПРИАМ) // Каталог программные продукты Украины. К.: 1993. С. 24-27.
4. *С.Н. Лапач* Проблемы построения математических моделей экспериментально-статистическими методами / С.Н. Лапач Проблемы построения математических моделей экспериментально-статистическими методами // Прогресивна техніка і технологія машинобудування, приладобудування і зварювального виробництва. Праці НТУУ "КПІ", –Т. 2, –К.: НТУУ "КПІ", –1998. -С.25-29.
5. *С.Н. Лапач* Статистические методы в медико-биологических исследованиях с использованием Excel / С.Н. Лапач, А.В. Чубенко, Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel – 2 изд. перераб. и доп.–К.: 2001, Морион. – 408с.
6. *Радченко С.Г.* Методология регрессионного анализа / Радченко С.Г. Методология регрессионного анализа –К.: «Корнійчук», 2011. –376с.
7. *Иванов Г.А.* Статистические методы восстановления истинной зависимости по экспериментальным данным / Иванов Г.А. Турбан А.Ф. Статистические методы восстановления истинной зависимости по экспериментальным данным.— К.: Знание.— 1986.— 22 с.

References

1. *Kolesov I.M.* Osnovy tehnologi mashinostroenija: Ucheb. dlja mashinostroit. spec. vuzov (Bases of technology of engineer: Studies. for mashinostroit. special. institutes of higher). Moskow: Vyssh. shk., 2001.591p.
2. *Kacev P.G.* Statisticheskie metody issledovanija rezhushhego instrumenta. Izd. 2-e, pererab. i dop.(Statistical methods of research of cutting tool. 2-nd ed.). Moskow: Mashinostroenie, 1974. 231p.
3. *Lapach S.N.,* Radchenko S.G., Babich P.N. Planirovanie, regressija i analiz modelej PRIAM (PRIAM). Katalog programmnye produkty Ukrainy. [Design, regression and analysis of models of PRIAM: Catalogue software products of Ukraine]. Kyiv.: 1993. pp. 24-27.
4. *S.N. Lapach* Problemy postroenija matematicheskih modelej jeksperimental'no-statisticheskimi metodami. Progresivna tehnika i tehnologija mashinobuduvannja, priladobuduvannja i zvarjuval'nogo virobnictva. Praci NTUU "KPI", T. 2, [Problems of construction of mathematical models statistical methods: A progressive technique and technology of engineer, instrument-making and welding production]. Kyiv: NTUU "KPI", 1998. pp.25-29.
5. *S.N. Lapach*, A.V. Chubenko, Babich P.N. Statisticheskie metody v mediko-biologicheskih issledovanijah s ispol'zovaniem Excel –2 izd. pererab. i dop.(Statistical methods in medics and biology researches with the use of Excel – 2nd Ed.). Kyiv: 2001, Morion. 408p.
6. *Radchenko S.G.* Metodologija regressionnogo analiza (Regression analysis methodology).Kyiv.: «Kornijchuk», 2011. 376p.
7. *Ivanov G.A.* Turban A.F. Statisticheskie metody vosstanovlenija istinnoj zavisimosti po jeksperimental'nym dannym (Statistical methods of reducing of true dependence on experimental data). Kyiv: Znanie. 1986. 22 p.

Подана до редакції 27.09.2014